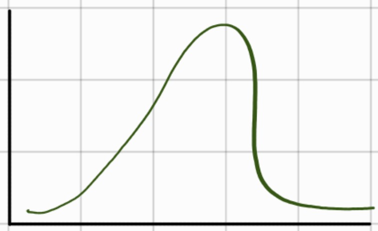
* Fréchet proposed a distribution separately for extreme values. Although the data is time dependent so it is a markov chain, we only extract the extreme values and fit a distribution on them.

* Fisher and Tippett published three asymptotic forms of extreme value distribution.

* Gumbel made significant contribution in EVT field.

* Extreme events are experienced in natural and engineering systems. EVT develops models and frameworks for describing the unusual events (quantify comparatively large / small undesired events).

* In classical statistics, we are bothered about the average behaviour of the stochastic process (CLT).



$X \sim \text{Unif}[a, b]$

EVT focuses on extreme and rare events

$X_{(n)} \sim \text{Unif}[a, b]$

Governed by Frechet-Fisher Tippett Gnedenko

We are modelling the tail behaviour of the distribution.

* Goals : Explore extreme events (emergence of extreme events)
Predicting extreme events
Design control strategies for mitigating extreme events.

* There are two approaches : 1) Block maxima (minima)
2) Peak over threshold

In Block maxima approach, $M_n = \{X_1, X_2, \dots, X_n\}$ for $n \to +\infty$
M_n follows a GEV distribution.

In Peak over threshold, $\{X_i - u \mid X_i > u\}$ and large u this follows a GPD (generalized pareto distribution).

* By CLT : $\dfrac{\sqrt{n}\,(\bar{X} - \mu)}{\sigma} \sim N(0,1)$

* EVT : $M_n = [F(x)]^n \longrightarrow \begin{cases} 0 , & F(x) < 1 \\ 1 , & F(x) = 1 \end{cases}$ degenerate function

$\mathbb{P}(X_1 \geq x, \dots, X_n \geq x)$

Can we find (similar to CLT) transformation to avoid degeneracy of M_n and find $(a_n), (b_n)$; $\lim_{n \to +\infty} \mathbb{P}\left(\frac{M_n - b_n}{a_n} \leq x\right) = H(x)$

$\mathbb{P}(M_n \leq x) = (F(x))^n$

$\mathbb{P}\left(\frac{M_n - b_n}{a_n} \leq x\right) = \mathbb{P}(M_n \leq b_n + a_n x) = (F(b_n + a_n x))^n = H(x)$

a non degenerate DF.

This gave birth to "three type" theorem.

* **Theorem** :

Let $\{X_i\}$ be a sequence of IID RVs. If there exists normalizing constants $b_n \in \mathbb{R}$, $a_n > 0$ and some non-degenerate function H; $\frac{M_n - b_n}{a_n} \xrightarrow{d} H$. Then, H belongs to the type of one of the following three DFs :

1) $H_{1,\alpha}(x) := \exp(-x^{-\alpha}) \, \mathbb{1}_{\{x>0\}}$, $\alpha > 0$ Fréchet

2) $H_{2,\alpha}(x) := \mathbb{1}_{\{x \geq 0\}} + \exp\{-(-x)^{\alpha}\} \, \mathbb{1}_{\{x<0\}}$ $\alpha > 0$ Weibull

3) $H_{3,0}(x) := \exp(-e^{-x}) \quad \forall x \in \mathbb{R}$ Gumbel

We call any distribution $F \in MDA(H) \Leftrightarrow \exists (a_n) > 0, b_n \quad \forall x \in \mathbb{R}$

$\downarrow$

Maximum domain of attraction $\qquad \lim_{n \to +\infty} F^n(a_n x + b_n) = H(x)$

* **Theorem** : Extreme Value Theorem

If $F \in MDA(G)$, then necessarily G is of the same type as the GEV CDF H_ξ $\left(G(x) = H_\xi(ax+b) \quad (a>0) \right)$

$$G_{\mu,\sigma,\xi}(x) = \begin{cases} \exp\left(-\left(1+\xi \frac{x-\mu}{\sigma}\right)^{-1/\xi}\right) & , \quad 1+\xi\frac{x-\mu}{\sigma} > 0 \\ \exp\left(-e^{-\left(\frac{x-\mu}{\sigma}\right)}\right) & , \quad \xi = 0 \end{cases}$$

μ : location

σ : scale

ξ : shape

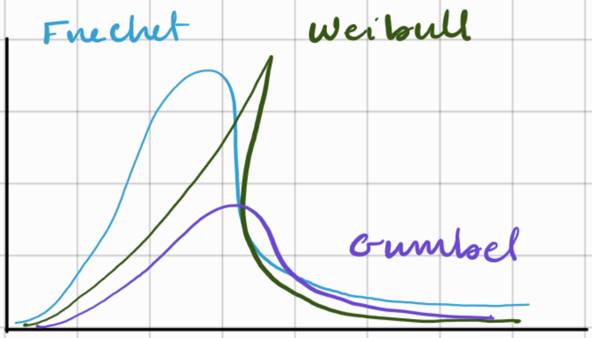
$\xi = \frac{1}{\alpha}$ determines the nature of the tail distribution, ξ is called the tail index.

Order of importance : Fréchet > Gumbel > Weibull (hardly used)

* $\xi > 0$: Fréchet (Heavy tailed distribution) polynomial decay

$\xi < 0$: Weibull (light tailed distribution) finite upper end

$\xi = 0$: Gumbel (exponential decay)



MDA	Distribution
Frechet	Pareto, Burr, Student t, cauchy
Weibul	Uniform, Beta, Kumaraswm
Gumbel	Exp, normal, Gamma

Disadvantage : Events that are not actually extreme points may be regarded as extremes when choosing one maximum per block.

Extreme events come in clusters in a short period in time so it cannot be captured in a single block.

* Peaks over threshold tries to solve the problems of BM.

$Y_i = \{ e_i - u \mid e_i > u \}$ will asymptotically follow a generalised pareto distribution. This is called the Pickand Theorem.

* Theorem : The Pickand Theorem

Let us consider a sequence of iid rvs $X_1, \ldots, X_n$ having CDF F. For a sufficiently large threshold u, $\exists\, \sigma_u > 0$ and ξ real number such that GPD is a good approximation of the excess $(X_i - u)$

$$F_u(y) := \mathbb{P}(X - u \leq y \mid X > u) \approx G_{\xi, \sigma_u}(y)$$

$$G_{\xi,\sigma_u}(y) = \begin{cases} 1 - \left(1 + \xi \dfrac{y}{\sigma_u}\right)^{-1/\xi}, & \xi \neq 0 \end{cases} \qquad y \geq 0$$

ξ : shape

$$\left(\; 1 - e^{-y/\sigma_u} \qquad , \quad \xi = 0 \right.$$ σ_u : scale

$\xi > 0$: $\bar{G}_{\xi,\sigma_u}(y) = 1 - G_{\xi,\sigma_u} \sim c \cdot y^{-1/\xi}$, $c > 0$ heavy tailed (Fréchet

$\xi = 0$: $\bar{G}_{\xi,\sigma_u}(y) = e^{-y/\sigma_u}$, light tailed / exponential (Gumbel)

$\xi < 0$: $G \in$ Weibull MDA

Disadvantage : Very sensitive to threshold

" How to select the threshold ? " Mean Excess Plot

- Mean Excess Plot Method :

If $Y \sim GPD_{\xi,\sigma_2}$ then its mean excess function (mean residual f" or life expectancy f") is linear ;

$$ME(u) = \mathbb{E}(Y - u \mid Y > u) = \frac{1}{1-\xi}(\sigma_u + \xi_u) \text{ where } \sigma_\xi + \xi_u > 0$$

Hence, above u at which the GPD provides a valid approximation to the excess distribution, the MEP should stay reasonably close to a linear function. This provides a way to select the threshold u.

A natural estimate of $ME(u)$ is the empirical MEF ($e(u)$)
$e(u) = \frac{1}{n_u} \sum_{i=1}^{n_u} (x_i - u)$; $u \geq 0$ where the data has n_u observations exceeding u. The empirical mean excess plot will converge in prob. to a straight line of the distribution n of excesses is fitted with a GPD with finite mean. Once u is chosen we can use MLE or MOM to find tail index ξ and that scaling parameter σ.

* Disadvantages : We choose a threshold after which the ME Plot becomes linear but this can only give us a few extreme points using this threshold. There are other plots that can be adopted for selecting the threshold rule of thumb, hill, qqplot.